PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Methods Used to Conduct and Report Bayesian Mixed Treatment Comparisons Published in the Medical Literature: A Systematic Review
AUTHORS	Sobieraj, Diana; Cappelleri, Joseph; Baker, William; Phung, Olivia; White, C; Coleman, Craig

VERSION 1 - REVIEW

REVIEWER	James (Jay) Brophy MEng MD FRCP(c) FACC PhD Professor of Medicine & Epidemiology (McGill University) Royal Victoria Hospital Montreal (Qc) CANADA
REVIEW RETURNED	28-Apr-2013

GENERAL COMMENTS	Overall I found the paper well written and of interest. I do have several comments for the authors to consider. 1. The objective of this manuscript is “To identify published Bayesian mixed treatment comparisons (MTCs) and to summarize characteristics regarding their conduction and reporting’. Given this objective, I am surprised by the authors’ decision to spend time describing Journal characteristics including impact factor, limitations on word counts, etc. The point of this data collection is not obvious to this reviewer. 2. I am not totally unsympathetic to the authors’ objectives, although the purely descriptive nature of the paper does dampen my enthusiasm. No hypotheses are presented and there is no attempt to determine what factors may correlate with higher quality reporting. The authors do recognize that a major limitation is that their observation of lack of reporting details does not mean these elements were not preformed. 3. However what I find a more interesting aspects of their data collection is not even discussed by the authors in the discussion Specifically, I strongly suspect that the quantity of evidence (number of studies included 36 and intervention 8-9) evaluated in MTC greatly exceeds that evaluated in standard meta-analyses (this could measured for one year to confirm his impression) and is worthy of some comment. 4. Now I also appreciate that data collection will end before the paper is submitted, it is a weakness that it is now almost 2 years before this review process (July 2011). I would imagine that the manuscript has already been submitted to one or several other journals and it would have been appropriate to attempt some updating of the data. I presume that comments from previous reviewers have been addressed but some additional transparency on this question would be helpful. 5. The methods state MEDLINE was searched and then mentions

	American College of Physicians Journal Club, is this not included in MEDLINE. Why the special attention to that particular Journal? Figure 1 sates 5 articles form other sources. What sources? 6. What is the validity (sensitivity and specificity) of the authors' classification of methodologist? There are approximately 100 individuals in my Department of Epidemiology and Biostatistics and no more than 50% who seriously self identify as methodologists. This suggest that their classification scheme may be no better than a coin toss. 7. The discussion section could be much more tightly written. I would suggest the following simple format Paragraph 1 – summary of the their results Paragraph #2 – where the results fit with previous work; #3 limits #4 strengths #5 conclusions & future work.

REVIEWER	Daniel E. Jonas, MD, MPH Associate Professor, Division of General Medicine University of North Carolina--Chapel Hill United States I have no competing interests.
REVIEW RETURNED	22-May-2013

THE STUDY	The discussion should be updated, considering recent literature.
RESULTS & CONCLUSIONS	The discussion is outdated and does not adequately consider recent literature to put the findings into context.
GENERAL COMMENTS	Major comments: 1. This is an interesting paper and it makes an important contribution to the literature on mixed treatment comparisons. The paper could increase awareness of a reporting problem and could help to improve appropriate reporting in future publications. 2. The Discussion could be improved and should be updated to put the findings into context better. Paragraphs 2, 3, and 4 seem particularly outdated and could be improved by incorporating a better reflection of current literature on the topic. For example, it states that there is "…limited guidance as to how to conduct and report a MTC…" and just provides one reference to a report written by the authors of this paper. There are several published papers providing relevant guidance that have come out over the past year or so. For example, ISPOR Task Force on Indirect Treatment Comparisons (two publications in Value in Health, both in 2001) and a full issue of Research Synthesis Methods (June 2012, Vol 3, issue 2) edited by Georgia Salanti. 3. Regarding the definition of MTC used (page 7), I think the authors have introduced an additional criteria, requiring one closed loop, that many systematic reviewers and analysts would disagree with. I suggest the authors provide additional references that would support the definition of MTC that they have developed. I'm not convinced that one closed loop should be required for something to be considered an MTC. It is certainly required that both direct and indirect evidence be used, but there are other network

	patterns that can utilize both direct and indirect evidence (although not for the same comparison). One closed loop does seem like the most obvious way to use both direct and indirect comparisons within a single analysis, but why wouldn't a "ladder" network also be considered an MTC. For example, if we want to know how drugs A, B, C, and D compare for a given condition, and if we have studies of drug A vs. drug B, and of drug B vs. placebo, and placebo vs. drug C, and drug C vs. drug D, then we have direct evidence (e.g., A vs. B and C vs. D) and we have some indirect evidence to use (e.g., comparing B vs. C via placebo)---if we then conduct a network meta-analysis of those studies to determine how A/B/C/D compare, why wouldn't that be considered an MTC? 4. The definition of methodologist seems fairly narrow and may be flawed. Can the authors provide any information to support the validity and reliability of their measure? 5. The paper would be improved by having a line editor review it. Many sentences are not written very well. <u>Minor Comments:</u> <u>Abstract:</u> 1. Objectives. Conduction should be changed to conduct. 2. Results. Line 40. "Methods used to …" Delete 'or'. 3. Results. "…although rarely done pictorially…". This part of the sentence doesn't quite make sense given how the earlier part is written. Ranking of interventions is not something that can be done pictorially. Maybe the *results of ranking* can be *displayed* in a figure/picture. 4. Conclusions. Here in the abstract, and elsewhere in the paper, the choice of "thought leaders" seems odd. It seems that analysts, statisticians, systematic reviewers, clinicians, etc. would all find it important to have clarity on appropriate methods and reporting. <u>Article summary:</u> 1. Article focus. Change conduction to conduct. 2. Strengths and limitations. Bullet 3. Suggest changing "…we evaluated additional…" to "we evaluated reporting of additional…" <u>Introduction:</u> 1. Paragraph 3. As in my general comments, I haven't seen other sources require a closed loop to be present for something to be considered a MTC meta-analysis. Including both direct and indirect comparisons within a network meta-analysis (but not necessarily having both direct and indirect evidence *for the same comparison*) is typically the requirement to be called a MTC. Further explanation, support, or references should be provided for why the authors have

	included more narrow criteria in the definition, requiring a closed loop. Methods: 1. pg 8. Line 33. Change "..we as investigators determined…" to "we determined" Results: 1. 2nd paragraph. Suggest deleting last sentence ("The remaining analyses…"

VERSION 1 – AUTHOR RESPONSE

James (Jay) Brophy MEng MD FRCP(c) FACC PhD
Professor of Medicine & Epidemiology (McGill University)
Royal Victoria Hospital
Montreal (Qc) CANADA

1. The objective of this manuscript is "To identify published Bayesian mixed treatment comparisons (MTCs) and to summarize characteristics regarding their conduction and reporting'. Given this objective, I am surprised by the authors' decision to spend time describing Journal characteristics including impact factor, limitations on word counts, etc. The point of this data collection is not obvious to this reviewer.

As this manuscript stems from a commissioned AHRQ (federal) report, some of our objectives were dictated by the agency. We believe that our intent to look at these journal characteristics are reflected in the "reporting" portion of our objective. We also believe that these data are important since it is possible journal characteristics such as limitations and word counts could impact how a MTC was reported (that is, maybe a method was not described in the methods to shorten the manuscript, etc.). Moreover, we thought it would be interesting to understand more about where these analyses are published and to summarize the characteristics of those journals- a noteworthy endeavor in its own right.

2. I am not totally unsympathetic to the authors' objectives, although the purely descriptive nature of the paper does dampen my enthusiasm. No hypotheses are presented and there is no attempt to determine what factors may correlate with higher quality reporting. The authors do recognize that a major limitation is that their observation of lack of reporting details does not mean these elements were not preformed.

Thank you for this comment. It was our intention to present a descriptive report on this topic. It was not our objective to assess the quality of the included publications. Our research, though, may help to pave the way for future research with hypotheses.

3. However what I find a more interesting aspects of their data collection is not even discussed by the authors in the discussion Specifically, I strongly suspect that the quantity of evidence (number of studies included 36 and intervention 8-9) evaluated in MTC greatly exceeds that evaluated in standard meta-analyses (this could measured for one year to confirm his impression) and is worthy of some comment.

We agree with this comment and have added evidence to support the concept that MTCs are larger and more comprehensive that traditional meta-analyses by adding the following (with a reference) to the first paragraph of the discussion: “A recent study found that a median of 3 studies (interquartile range 2 to 6) were included per meta-analysis, with close to 75% of meta-analyses including five or less trials. Our results suggest that compared to traditional meta-analyses, closed-loop Bayesian MTCs are larger and more comprehensive analyses.”

4. Now I also appreciate that data collection will end before the paper is submitted, it is a weakness that it is now almost 2 years before this review process (July 2011). I would imagine that the manuscript has already been submitted to one or several other journals and it would have been appropriate to attempt some updating of the data. I presume that comments from previous reviewers have been addressed but some additional transparency on this question would be helpful.

Thank you for this comment. This is the first journal which we have submitted this manuscript for consideration. The lag time is due to the fact that this manuscript is reflecting a portion of an AHRQ report and we are not allowed to pursue publication of a manuscript until the full AHRQ report is published. It is not unusual for there to be a long lag time from the time of project completion, to full AHRQ report publication, and then, finally, to manuscript publication. AHRQ requires a manuscript reflect the same data/analysis/results as the published report before they will give clearance for publication, therefore, we have not updated these analyses.

5. The methods state MEDLINE was searched and then mentions American College of Physicians Journal Club, is this not included in MEDLINE. Why the special attention to that particular Journal? Figure 1 sates 5 articles form other sources. What sources?

The reviewer is correct about the ACP journal club. This source was not specifically targeted for special attention; however, the OVID platform used by our institution packages it along with the other non-MEDLINE databases we used. Thus to be fully accurate we report that we did a separate search of this database.

As for other sources, we allowed manual additions of full text publications based on the literature search results. These would typically have been identified as a review, published only in abstract form in the Centre for Reviews and Dissemination Databases, of a full text study that met criteria. Therefore we had to identify the full text citation the review addressed and manually add it. .

6. What is the validity (sensitivity and specificity) of the authors’ classification of methodologist? There are approximately 100 individuals in my Department of Epidemiology and Biostatistics and no more than 50% who seriously self identify as methodologists. This suggests that their classification scheme may be no better than a coin toss.

We concur that the definition used for “methodologist” has limitations and for that reason we have added this as a limitation to our review, in the discussion section. Of note, we did not ourselves create the definition; it has been previously used by a research group reviewing similar data.

7. The discussion section could be much more tightly written. I would suggest the following simple format Paragraph 1 – summary of the their results Paragraph #2 – where the results fit with previous work; #3 limits #4 strengths #5 conclusions & future work.

Thank you for this suggestion. We have now more closely followed the sequence suggested by this reviewer. We believe beginning with the importance of MTC in medical literature adds value to the work we have done and presented. We follow with a brief overview of our results without being overly repetitive given the descriptive nature of our review. We believe that prior to describing what our

review adds to current literature, a brief review of what has been previously done is essential and is the order which we follow. The remaining sections are consistent with the suggested sequence.

Reviewer: Daniel E. Jonas, MD, MPH
Associate Professor, Division of General Medicine
University of North Carolina--Chapel Hill
United States

Major comments:
1. This is an interesting paper and it makes an important contribution to the literature on mixed treatment comparisons. The paper could increase awareness of a reporting problem and could help to improve appropriate reporting in future publications.

Thank you.

2. The Discussion could be improved and should be updated to put the findings into context better. Paragraphs 2, 3, and 4 seem particularly outdated and could be improved by incorporating a better reflection of current literature on the topic. For example, it states that there is "…limited guidance as to how to conduct and report a MTC…" and just provides one reference to a report written by the authors of this paper. There are several published papers providing relevant guidance that have come out over the past year or so. For example, ISPOR Task Force on Indirect Treatment Comparisons (two publications in Value in Health, both in 2001) and a full issue of Research Synthesis Methods (June 2012, Vol 3, issue 2) edited by Georgia Salanti.

Thank you for this comment. This manuscript reflects one of three parts to the AHRQ commissioned report we published on The Use of Mixed Treatment Comparison in Systematic Review found here http://www.effectivehealthcare.ahrq.gov/ehc/products/354/1238/Use-of-Mixed-Treatment_FinalReport_20120823.pdf The first part of this report summarizes all available guidance documents and includes all of the references this reviewer has listed and so we believe the reader would gain a comprehensive review of all available guidance by referring to our full report, rather than just selectively listing a few citations in this manuscript. We have revised the sentence (which was moved to comply with Reviewer 1 comment #7) to make it clear this has been reviewed and summarized elsewhere to encourage the reader, if they are interested, in accessing the review "This may be related to the limited guidance as to how to conduct and report a MTC, a topic which has been extensively reviewed and summarized elsewhere.[11]

We also cite the ISPOR documents (references 3 and 4) earlier in our manuscript as well as the publication by Salanti et al (reference 12).

3. Regarding the definition of MTC used (page 7), I think the authors have introduced an additional criteria, requiring one closed loop, that many systematic reviewers and analysts would disagree with. I suggest the authors provide additional references that would support the definition of MTC that they have developed. I'm not convinced that one closed loop should be required for something to be considered an MTC. It is certainly required that both direct and indirect evidence be used, but there are other network patterns that can utilize both direct and indirect evidence (although not for the same comparison). One closed loop does seem like the most obvious way to use both direct and indirect comparisons within a single analysis, but why wouldn't a "ladder" network also be considered an MTC. For example, if we want to know how drugs A, B, C, and D compare for a given condition, and if we have studies of drug A vs. drug B, and of drug B vs. placebo, and placebo vs. drug C, and drug C vs. drug D, then we have direct evidence (e.g., A vs. B and C vs. D) and we have some indirect evidence to use (e.g., comparing B vs. C via placebo)---if we then conduct a network meta-analysis of those studies to determine how A/B/C/D compare, why wouldn't that be considered an MTC?

This is a good point and we agree. It was not our intent to state that MTCs must have a closed loop, but rather that our analysis (per the direction of AHRQ) was focused/restricted to the situation of MTCs with at least one closed loop. We have revised the text throughout the manuscript to clarify this.

4. The definition of methodologist seems fairly narrow and may be flawed. Can the authors provide any information to support the validity and reliability of their measure?

We concur that the definition used for “methodologist” has limitations and for that reason we have added this as a limitation to our review, in the discussion section. Of note, we did not ourselves create the definition; it has been previously used by a research group reviewing similar data.

5. The paper would be improved by having a line editor review it. Many sentences are not written very well.

Thank you for this comment. Three new individuals have reviewed and edited the manuscript for clarity and language and, as a result, we believe it reads better.

Minor Comments:
Abstract:
1. Objectives. Conduction should be changed to conduct.

This has been changed as suggested.

2. Results. Line 40. “Methods used to …” Delete ‘or’.

This has been changed as suggested.

3. Results. “…although rarely done pictorially…”. This part of the sentence doesn’t quite make sense given how the earlier part is written. Ranking of interventions is not something that can be done pictorially. Maybe the results of ranking can be displayed in a figure/picture.

We agree with the comment and have changed “done” to “displayed”.

4. Conclusions. Here in the abstract, and elsewhere in the paper, the choice of “thought leaders” seems odd. It seems that analysts, statisticians, systematic reviewers, clinicians, etc. would all find it important to have clarity on appropriate methods and reporting.

We agree with the comment and thought that our use of “thought leaders” summarized the individuals listed, although we are happy to delete “thought leaders” so as not to imply one group over another.

Article summary:

1. Article focus. Change conduction to conduct.

We have made this change as suggested.

2. Strengths and limitations. Bullet 3. Suggest changing “…we evaluated additional…” to “we evaluated reporting of additional…”

We have made this change as suggested.

Introduction:

1. Paragraph 3. As in my general comments, I haven't seen other sources require a closed loop to be present for something to be considered a MTC meta-analysis. Including both direct and indirect comparisons within a network meta-analysis (but not necessarily having both direct and indirect evidence for the same comparison) is typically the requirement to be called a MTC. Further explanation, support, or references should be provided for why the authors have included more narrow criteria in the definition, requiring a closed loop.

This is a good point and we agree. It was not our intent to state that MTCs must have a closed loop, but rather that our analysis (per the direction of AHRQ) was focused/restricted to the situation of MTCs with at least one closed loop. We have revised the text throughout the manuscript to clarify this.

Methods:

1. pg 8. Line 33. Change "..we as investigators determined…" to "we determined"

We have made this change as suggested.

Results:

1. 2nd paragraph. Suggest deleting last sentence ("The remaining analyses…"

We have deleted this sentence as suggested.

VERSION 2 – REVIEW

REVIEWER	Brophy, james McGill University
REVIEW RETURNED	16-Jun-2013

GENERAL COMMENTS	The authors have addressed all my previous remarks and concerns. I would recommend acceptance.